## (12) EUROPEAN PATENT APPLICATION

(72) Inventor : Chen, Homer H,
269 Harvey Avenue
Lincroft, New Jersey 07738 (US)
Inventor : Civanlar, Mehmet R.
85 Coleman Avenue
Red Bank, New Jersey 07701 (US)
Inventor : Haskell, Barin G.
82 Glenwood Drive
Tinton Falls, New Jersey 07724 (US)

(74) Representative : Buckley, Christopher Simon
Thirsk et al
AT&T (UK) LTD.,
AT&T Intellectual Property Division,
5 Mornington Road
Woodford Green, Essex IG8 0TU (GB)

(54) Block transform coder for arbitrarily shaped image segments.

(57) A Transform Coder Unit (TCU) (6) for trans-
forming an arbitrarily shaped image into opti-
mal transform coefficients (OTC) for data
transmission comprises a forward transform (7)
which transforms the image to transform coeffi-
cients, and a TCS generator (8) which generates
a transform coefficient set (TCS) from the trans-
form coefficients. The TCU also contains an
inverse transform (9) which transforms the TCS
to a computed region block having computed
pel values. Finally, the TCU comprises a rep-
lacer (10) which replaces those computed pel
values corresponding to the interior pel set with
the original pel values to form a modified com-
puted region block which is re-iterated until
optimal transform coefficients are determined.
    A process for determining optimal transform
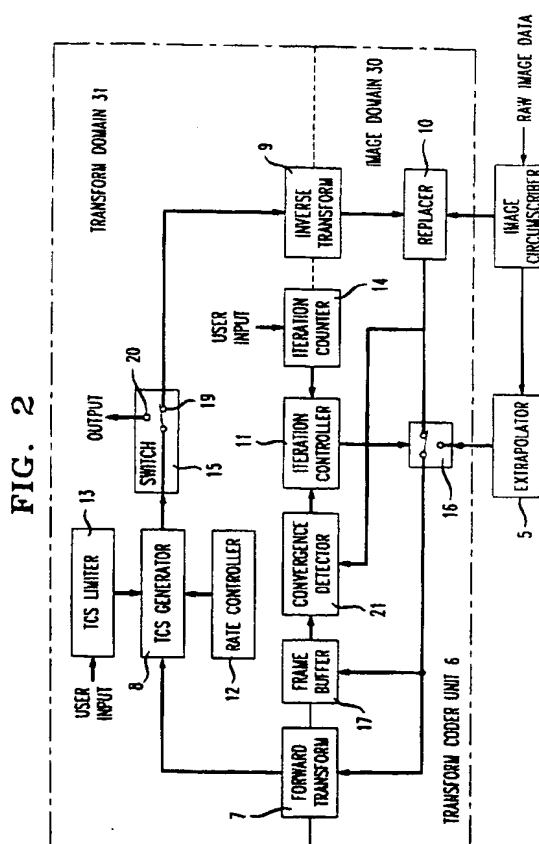coefficients using the aforementioned device is
described.

FIG. 2

EP 0 649 258 A2

## Field of the Invention

The present invention relates generally to a method and device to code images for data transmission, and more specifically to a method and device to determine the optimal transform coefficients for an irregular shaped image for low bit-rate transmission using standard transforms.

## Information Disclosure Statement

Although current video coding standards may operate at very low bitrates, the trade-off between temporal and spatial resolution results in visually annoying motion or spatial artifacts. Therefore, the International Organization for Standardization is considering developing a new standard for very low bitrate A/V coding. ISO/IEC JTC1/SC29/WG11 MPEG 92/699, "Project Description for Very- Low Bitrate A/V Coding" (Nov. 5, 1992). This document reviews the state of the art and proposes a direction for future research.

In typical image coding systems, the image to be coded is usually processed using $N \times N$ blocks of picture elements (pels) regardless of the image content. This approach, however, may lead to visible distortions known as blocking and mosquito effects, particularly at low bit-rates. To avoid these visual artifacts, region-based image representation partitions the image into regions of similar motion or texture, yielding image segments of arbitrary shape instead of fixed (rectangular) blocks. Such image representation offers several advantages over the conventional block-based representation such as adaptation to local image characteristics. Consequently, region-based image representation has received considerable attention in MPEG4 video coding standard work for very low bitrate coding.

A fundamental issue in region-based image compression is the coding of arbitrarily shaped image segments. An arbitrarily shaped image segment f(x,y) can be approximated by a set of basis functions optimized for the shape of the image segment to be coded:

$$f(x,y) = \sum_i \alpha_i \phi_i(x,y) \qquad (1)$$

where x,y $\varepsilon$ S, S is the region occupied by the image segment, $f(x,y)$ is the approximation of the image segment, and $\phi_i$'s are the basis functions. However, such shape-adapted transform techniques require a large amount of memory for storing the set of basis functions. As a result, these techniques are only suitable for small regions. Furthermore, for each new segment a new set of basis functions has to be computed. Thus, extensive computation is involved. Since no fast algorithms exist, these techniques are not attractive for practical use.

Another popular approach is to use one of the most popular image compression techniques, transform coding. In transform coding, an image is transformed from the image intensity domain to a new domain prior to coding and transmission. The new domain is selected so that the energy of the image becomes concentrated to a small region in the new domain. Among the various transforms, the discrete cosine transform (DCT) is the most widely used transform. It has become the industry standard because it provides a good approximation of the optimal Karhunen-Loeve transform (KLT) for a certain class of images, and can be computed by means of fast algorithms.

With block transform coding, the image segment can be approximated by a set of two-dimensional basis functions defined on a rectangular block "B" which circumscribes the image:

$$f(x,y) = \sum_i \beta_i \psi_i(x,y) \qquad (2)$$

where x,y $\in$ S, and $\psi_i$'s are the basis functions defined on the full block 3. The best approximation $f(x,y)$ of an image segment can be found by minimizing the squared error between the image segment and the approximation, i.e.,

$$\text{error} = \Sigma (f(x,y) - f(x,y))^2 \qquad (3)$$

This is equivalent to solving the Gaussian normal equations. Note that the summation is taken over the region defined by the image segment; pels outside the region are discarded. Since the number of pels of the image segment is usually less than the number of basis functions, the problem is undetermined, and several solutions are possible. To arrive at a single solution, the problem can be solved by successive approximation. This involves starting with a small subset of basis functions and exhaustively searching for the best solution. Although successive progression will yield a solution, the computational cost is high. Furthermore, like the shape-adapted techniques, no fast algorithms are available to make real-time implementation possible.

A more efficient approach is to perform the transform on the entire block,

$$f(x,y,) = \sum_i \gamma_i \psi_i(x,y) \qquad\qquad (4)$$

5　where $x, y \in B$, and $B$ is the area of the block. The transform can be performed in real-time by special purpose chips designed for block transforms. However, this technique requires that the pels outside the image segment be initialized before the transform occurs. The outside pels can be chosen such that the sum of squared errors over the image segment expresses by Equation (3) is minimized. This approach enables the transform spectrum to be optimized by choosing appropriate pel values outside the image segment. To this end, zeroing the
10　outside pels would be an easy way to initialize them. This approach, however, introduces discontinuities at the boundary of the image segment, yielding high frequency components that degrade the coding performance. To alleviate the problem, the image segments can be extrapolated outside the boundary by mirroring or pel repetition such that a smoother transformation can be obtained. This ad hoc approach though, fails to provide consistent, satisfactory results. Consequently, a more promising method is needed. The present in-
15　vention fulfills this need.

　　The present invention utilizes the theory of successive projection onto convex sets (POCS). In Patrick L. Combettes, "The Foundation of Set Theoretic Estimation," Proceedings of the IEEE, Vol. 81, No. 2 (Feb. 1993), this theory is described in a theoretical sense. The present invention applies this theory in a practical sense to image coding.

20

## Summary of The Invention

　　The present invention is directed at a method and a device for determining the optimal transform coefficients for an arbitrarily shaped image for data transmission. The invention uses block transforms with frequen-
25　cy domain region-zeroing and space domain region-enforcing operations for effectively coding arbitrarily shaped image segments. The block transform is computed over a rectangular block which circumscribes the arbitrary shape. To find the best values for a group of selected transform coefficients, the invention uses an iterative technique based on the theory of successive projection onto convex sets (POCS). A key feature of the technique is that it works with existing block transform coding hardware (such as DCT chips) and software.
30　Therefore, it can be implemented using existing codec component at an insignificant cost.

## Brief Description of The Drawings

　　Fig. 1 depicts an arbitrary shape and the circumscribed rectangular region.
35　　Fig. 2 shows a preferred embodiment of the TCU which detects convergence in the image domain.
　　Fig. 3 shows another preferred embodiment of the TCU which detects convergence in the transform domain.
　　Fig. 4 shows another preferred embodiment of the present invention wherein a multiplicity of TCU are connected in series.

40

## Detailed Description of The Present Invention

　　The present invention relates to an iterative technique to determine optimal transform coefficient values for the coding of arbitrarily shaped images. The convergence of the iteration to the optimal solution is guar-
45　anteed by the theory of successive projection onto convex sets (POCS). The technique can be described within the POCS context by using two sets of images.
　　The first set is defined based on a basic premise of transform coding -- the energy compaction property of transform coefficients. This property provides that a large amount of energy is concentrated in a small fraction of the transform coefficients, and only these coefficients need to be kept for coding the image. The set
50　of images which can be represented using a selected group of transform coefficients constitute the first set and will be referred to as the transform coefficients set (TCS). This set is convex for all linear and some non-linear transformations. The projection of an arbitrarily shaped image block onto this set can be determined by computing the block transform and selecting and retaining high energy coefficients. The remaining, non-selected coefficients are zeroed (region-zeroing in the frequency domain).
55　　The second set is derived form the fact that the values of the pels outside of the arbitrary shaped region are irrelevant to coding. Thus, the second set becomes the set of images whose pel values within the arbitrarily shaped region are specified by the image to be coded. This set is referred to as the region of support set (RSS). This set is convex. The projection of an arbitrarily shaped region onto this set can be obtained by replacing

3

those pel values corresponding to the image's interior pels with the original pel values (region-enforcing in the space domain). This theory provides the basis for the present invention.

The present invention basically comprises two parts. Fig. 1 depicts the first part which involves generating and preparing the data to be coded. In this step, a rectangular region block is circumscribed around an arbitrarily shaped image 2. This defines an original internal pel set 3 which lies within arbitrarily shaped image 2 and within region block 1, and an original external pel set 4 which lies outside arbitrarily shaped image 2 and within region block 1.

To initialize the pel values of external pel set 4, an extrapolator 5 extrapolates the pel values of internal pel set 3. Examples of extrapolation methods include mirroring or pel repetition of the segments of internal pel set 3. Once external pel set 4 is initialized, the image data can be manipulated in the second part.

The second part involves a transform coder unit (TCU) 6 performing a POCS iteration loop on the image data. TCU 6 is shown in Fig. 2. TCU 6 comprises a forward transform 7, which operates at real-time and transforms the image from the image domain 30 to the transform domain 31.

Next, a TCS generator 8 generates a transform coefficient set (TCS) from the transform coefficients. This can be accomplished in a couple of ways. First, TCS generator 8 may contain a quantizer which generates the TCS by quantizing the transform coefficients. There is no convergence guarantee, however, under this alternative. A more preferred embodiment utilizes the energy compaction property of transform coefficients. This property holds that a large amount of energy is concentrated in a small fraction of the transform coefficients. Therefore, TCS generator 8 need only select and retain these coefficients for coding the image. The remaining transform coefficients can be zeroed.

If the energy compaction property is used to generate the TCS, then the number of coefficients to retain should be established. This may accomplished via a rate controller 12. Rate controller 12 can establish the threshold energy level at which to retain coefficients based on the size of the arbitrarily shaped image, and the bit budget of the encoder which will eventually code the transform coefficients. Alternatively, the number of transform coefficients to retain can be established independently via a TCS limiter 13 at the beginning of each iteration. A combination of both these mechanisms could be used as well.

TCS generator 8 outputs the TCS from the TCU if the TCS represents the optimal transform coefficients (OTC). Otherwise, TCS generator 8 sends the TCS to an inverse transform 9. Inverse transform 9 converts the TCS from transform domain 31 to image domain 30, thereby producing a computed regional block having computed pel values.

A replacer 10 replaces those computed pel values corresponding with internal pel set 3 with the original pel values, thereby forming a modified computed regional block (MCRB). The MCRB is then re-iterated through a reiterative forward transform. In the preferred embodiment of Figs. 2 and 3, the re-iterative forward transform and forward transform 7 are the same. Thus, the same TCU will re-iterate the MCRB.

The re-iterative forward transform and forward transform 7, however, can be different. For example, Fig. 4 shows a successive connection of TCUs 201-204. In this configuration, the re-iterative forward transform of TCU 201 is the forward transform of succeeding TCU 202. Thus, the modified computed region block is re-iterated through different TCUs. The number of TCUs in series determines the number of iterations performed.

Although the number of iterations depends upon the number of successive TCUs in the embodiment of Fig. 4, the number of iterations is variable in the embodiments of Figs. 2 and 3. Consequently, an iteration controller 11 is employed in both embodiments. Referring only to Fig. 2, iteration controller 11 controls switch 15 which has a first position 19 and a second position 20. First position 19 directs the TCS from TCS generator 8 to inverse transform 9 when the TCS does not represent the OTC. Second position 20 directs the TCS from TCS generator 8 to a quantizer when the TCS represents the OTC.

Iteration controller 11 may control the switching of switch 15 through a couple of mechanisms. As Fig. 2 shows, an iteration counter 14 can be used to count the number of iterations. When a pre-determined number is reached, iteration counter 14 will signal iteration controller 11 which will move switch 15 from first position 19 to second position 20.

Fig. 2 depicts another method of controlling switch 15 by monitoring image domain 30 of the TCU. Here, a convergence detector 21, and a frame buffer 17 are employed. Frame buffer 17 stores the pel values of the previous iteration. Convergence detector 21 switches switch 15 from first position 19 to second position 20 when the mean squared difference between the computed pel values stored in frame buffer 17 and those of the current iteration reaches a pre-determined level.

Fig. 3 depicts a device which also controls switch 115, but does so by monitoring transform domain 131 of TCU 106 using a convergence detector 121, and a frame buffer 117. Frame buffer 117 stores the TCS of the previous iteration. Convergence detector 121 switches switch 115 from first position 119 to second position 120 when the mean squared difference between the TCS stored in frame buffer 117 and that of the current iteration reaches a pre-determined level.

4

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.

**Claims**

1. A method of transforming an arbitrarily shaped image into optimal transform coefficients (OTC) for data transmission, said method comprises the steps of:
   a. generating original pel values by:
      (i) circumscribing said arbitrarily shaped image with a rectangular region block, thereby creating an internal pel set which lies within said arbitrarily shaped image and within said region block, and an external pel set which lies outside said arbitrarily shaped image and within said region block; and,
      (ii) initializing pel values of said external pel set by extrapolating the pel values of said internal pel set; and
   b. operating a transform coder unit (TCU) which calculates optimal transform coefficients by:
      (i) performing a forward transform on said region block to generate transform coefficients;
      (ii) generating a transform coefficient set (TCS) from said transform coefficients;
      (iii) performing an inverse transform on said TCS thereby generating a computed region block having computed pel values;
      (iv) replacing those computed pel values corresponding to said internal pel set with original pel values to form a modified computed region block (MCRB);
      (v) determining whether said TCS represents said OTC;
      (vi) reiterating steps (i) and (ii) on said modified computed region block and outputing said TCS when said TCS represents OTC; and,
      (vii) reiterating steps (i) through (vii) on said modified computed region block when said TCS values do not represent said OTC.

2. The method of claim 1.b. (i) wherein said forward transform uses a discrete cosine transform (DCT) chip.

3. The method of claim 1.b. (ii) wherein generating said TCS comprises quantizing said transform coefficients.

4. The method of claim 1.b. (ii) wherein generating said TCS comprises selecting and retaining those transform coefficients which have high energy according to the energy compaction property of transform coefficients, and zeroing the non-selected transform coefficients

5. The method of claim 4 wherein selecting said TCS comprises using a rate controller to establish a threshold energy level at which transform coefficients are retained, said rate controller establishes said level based on the bit budget of an encoder and the size of said arbitrarily shaped image.

6. The method of claim 4, wherein selecting said TCS comprises independently establishing a number of said transform coefficients to retain.

7. The method of claim 1.b. (v), wherein determining whether said TCS represents said OTC comprises independently establishing the number of iterations to perform.

8. The method of claim 1.b. (v), wherein determining whether said TCS represents said OTC comprises calculating when the mean squared difference between said MCRB of one iteration and that of a subsequent iteration reaches a pre-determined threshold.

9. A method of transforming an arbitrarily shaped image into optimal transform coefficients (OTC) for data transmission, said method comprises the steps of:
   a. generating original pel values by:
      (i) circumscribing said arbitrarily shaped image with a rectangular region block, thereby creating an internal pel set which lies within said arbitrarily shaped image and within said region block, and an external pel set which lies outside said arbitrarily shaped image and within said region block; and,
      (ii) initializing pel values of said external pel set by extrapolating the pel values of said internal pel set; and

5

b. operating a transform coder unit (TCU) to calculate optimal transform coefficients by:

(i) performing a forward transform on said region block to generate transform coefficients;

(ii) generating a transform coefficient set (TCS) from said transform coefficients;

(iii) determining whether said TCS represents optimal transform coefficients (OTC);

(iv) outputing said TCS when said TCS represents said OTC;

(v) performing an inverse transform on said TCS when said TCS does not represent said OTC, said inverse transform generates a computed region block having computed pel values;

(vi) replacing those computed pel values corresponding to said internal pel set with original pel values to form a modified computed region block; and,

(vii) reiterating steps (i) through (vii) on said modified computed region block.

10. A Transform Coder Unit (TCU) to transform an arbitrarily shaped image into optimal transform coefficients (OTC) for data transmission, said arbitrarily shaped image having original pel values, an interior pel set which lies within said image, and an exterior pel set which lies outside said image and within a rectangular region circumscribing said image, said TCU comprising:

a. a forward transform which transforms said image to transform coefficients;

b. a TCS generator which generates a transform coefficient set (TCS) from said transform coefficients, said TCS generator outputs said TCS when said TCS represents said OTC, and sends said TCS to an inverse transform when said TCS does not represent said OTC;

c. an inverse transform which transforms said TCS to a computed region block having computed pel values; and

d. a replacer which replaces those computed pel values corresponding to said interior pel set with said original pel values to form a modified computed region block (MCRB), said replacer sends modified computed region block to a re-iterative forward transform for re-iteration.

11. The TCU of claim 10, wherein said TCS generator includes a quantizer which generates said TCS by quantizing said transform coefficients.

12. The TCU of claim 10, wherein said TCS generator generates said TCS by selecting and retaining those transform coefficients which have high energy according to the energy compaction property of transform coefficients, and by zeroing all the non-selected transform coefficients.

13. The TCU of claim 12 wherein said TCS generator comprises a rate controller to establish a threshold energy level at which said TCS selector retains transform coefficients, said rate controller establishes said level based on the bit budget of an encoder and the size of said arbitrarily shaped image.

14. The TCU of claim 12 wherein said TCS generator comprises a TCS limiter to independently establish the number of transform coefficients to retain.

15. The TCU of claim 10, wherein said re-iterative forward transform and said forward transform are one in the same, and further comprising:

e. an iteration controller which controls an iteration switch having a first position and a second position, said first position directs TCS from said TCS generator to said inverse transform when said TCS does not represent said OTC, said second position directs said TCS from said TCS generator to output of said TCU.

16. The TCU of claim 15, wherein said iteration controller comprises an iteration counter to independently establish the number of iterations to perform, after said TCU performs the established number of iterations, said switch switches to said second position.

17. The TCU of claim 15, wherein said iteration controller contains a convergence detector, and a frame buffer, said frame buffer stores the pel values of a previous iteration, said convergence detector switches said switch from said first position to said second position when the mean squared difference between said MCRB stored in said frame buffer and that of the current iteration reaches a pre-determined level.

18. The TCU of claim 15, wherein said iteration controller contains a convergence detector, and a frame buffer, said frame buffer stores the TCS of a previous iteration, said convergence detector switches said switch from said first position to said second position when the mean squared difference between the TCS stored in said frame buffer and that of the current iteration reaches a pre-determined level.

6

**19.** The TCU of claim 10 wherein said re-iterative forward transform comprises a forward transform of a succeeding TCU, said succeeding TCU connected in series with said TCU.

**20.** The TCU of claim 10 wherein said forward transform is a discrete cosine transform (DCT) chip.
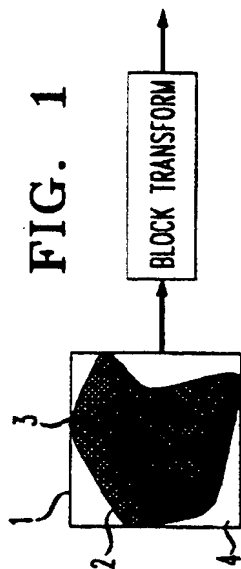
5

10

15

20
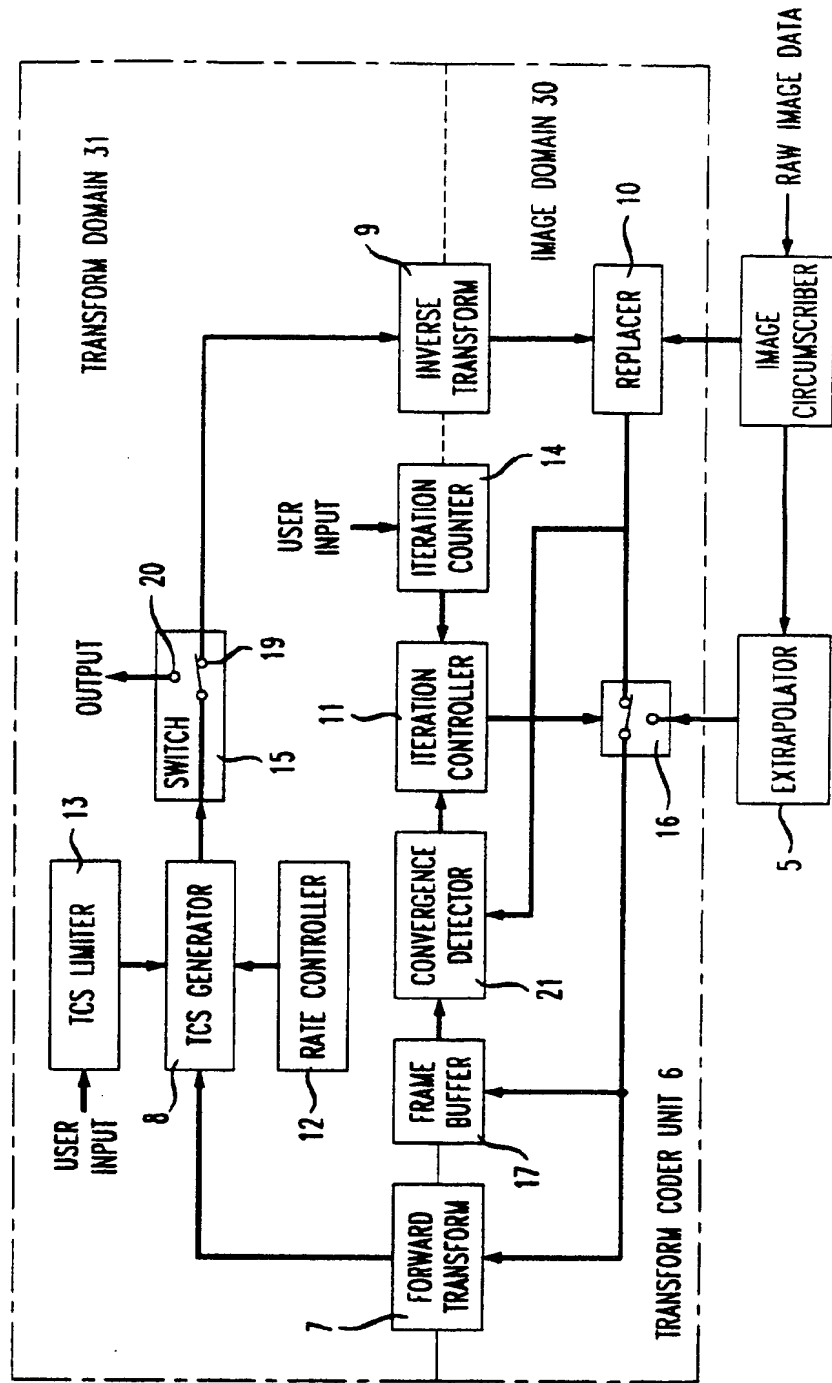
25

30

35

40

45

50

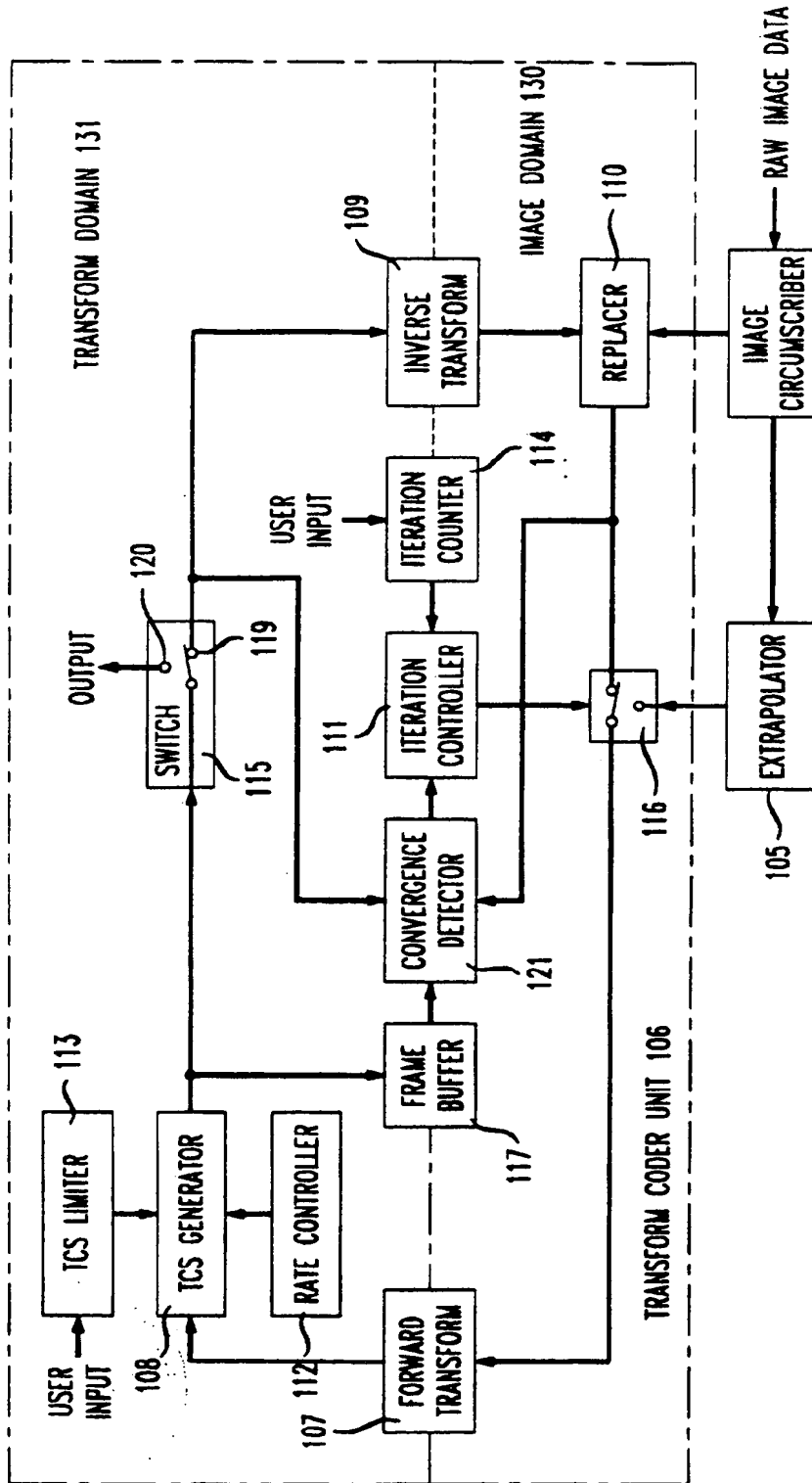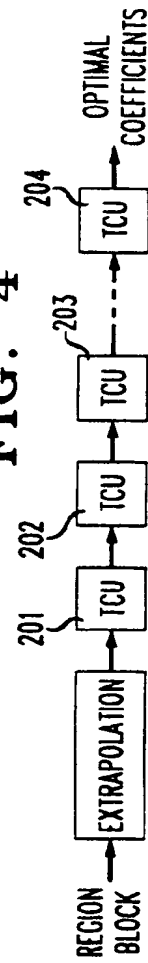55

# FIG. 1

BLOCK TRANSFORM

# FIG. 2

# FIG. 3



# FIG. 4



9

(12)                 **EUROPEAN PATENT APPLICATION**
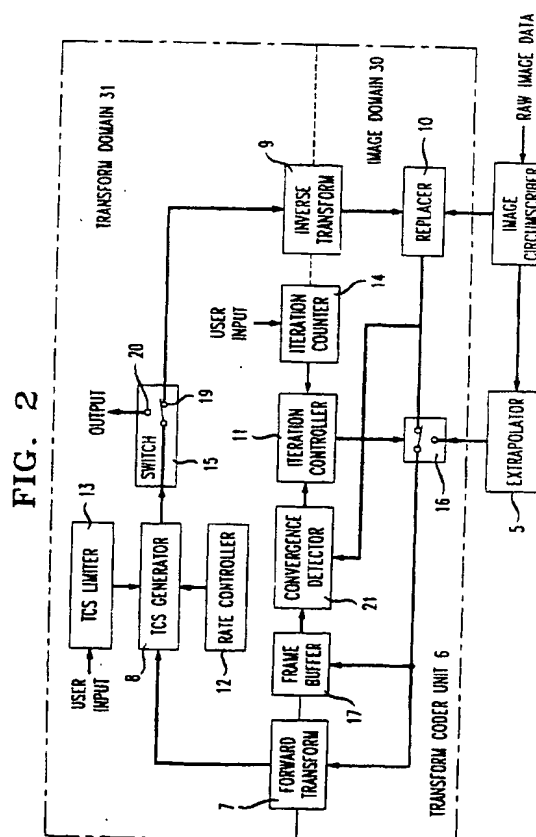
(72) Inventor : Chen, Homer H,
     269 Harvey Avenue
     Lincroft, New Jersey 07738 (US)
     Inventor : Civanlar, Mehmet R.
     85 Coleman Avenue
     Red Bank, New Jersey 07701 (US)
     Inventor : Haskell, Barin G.
     82 Glenwood Drive
     Tinton Falls, New Jersey 07724 (US)

(74) Representative : Buckley, Christopher Simon
     Thirsk et al
     AT&T (UK) LTD.,
     AT&T Intellectual Property Division,
     5 Mornington Road
     Woodford Green, Essex IG8 0TU (GB)

(54) **Block transform coder for arbitrarily shaped image segments.**

(57)    A Transform Coder Unit (TCU) (6) for trans-
forming an arbitrarily shaped image into opti-
mal transform coefficients (OTC) for data
transmission comprises a forward transform (7)
which transforms the image to transform coeffi-
cients, and a TCS generator (8) which generates
a transform coefficient set (TCS) from the trans-
form coefficients. The TCU also contains an
inverse transform (9) which transforms the TCS
to a computed region block having computed
pel values. Finally, the TCU comprises a rep-
lacer (10) which replaces those computed pel
values corresponding to the interior pel set with
the original pel values to form a modified com-
puted region block which is re-iterated until
optimal transform coefficients are determined.
    A process for determining optimal transform
coefficients using the aforementioned device is
described.

FIG. 2

EP 0 649 258 A3

## EUROPEAN SEARCH REPORT

**European Patent Office**

| Application Number |
|---|
| EP 94 30 7273 |

### DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| X | SIGNAL PROCESSING IV: THEORIES AND APPLICATIONS, 1 September 1988, ELSEVIER SCIENCE (NL) pages 1145 - 1148 FRANKE ET AL. 'Constrained Iterative Restoration Techniques: A Powerful Tool in Region Oriented Texture Coding' * the whole document * | 1-20 | H04N7/24 H04N7/26 H04N7/30 |
| X | SIGNAL PROCESSING V: THEORIES AND APPLICATIONS, vol.II, 18 September 1990, ELSEVIER SCIENCE (NL) pages 901 - 904, XP365734 YU ET AL. 'Pictorial Transform Coding for Tessellating Arbitrary Shaped Regions' * the whole document * | 1-20 | |
| A | ICASSP 87, vol.3, 6 April 1987, DALLAS, TEXAS, USA pages 1300 - 1303 FRANKE 'Selective Deconvolution: a New Approach to Extrapolation and Spectral Analysis of Discrete Signals' * the whole document * | 1,9,10 | **TECHNICAL FIELDS SEARCHED** (Int.Cl.6) H04N |
| A | DE-A-41 36 636 (BOSCH GMBH) * the whole document * | 1-20 | |
| A | PROCEEDINGS ACM MULTIMEDIA 93, 1 August 1993, ANAHEIM, CAL, USA pages 83 - 90 CHANG ET AL. 'Transform Coding of Arbitrarily-Shaped Image Segments' * paragraph 1 -paragraph 2 * | 1,9,10 | |

-/--

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 24 March 1995 | Foglia, P |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

2

## EUROPEAN SEARCH REPORT

European Patent Office

**Application Number**

EP 94 30 7273

### DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| A | 25TH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS & COMPUTERS, vol.2, 4 November 1991, PACIFIA GROVE, CAL, USA pages 1205 - 1209, XP314537 SANEI ET AL. 'Adapting Transform Coding to Texturally Segmented Images' * paragraph 2.4 * | 1,9,10 | |

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 24 March 1995 | Foglia, P |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another
    document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or
    after the filing date
D : document cited in the application
L : document cited for other reasons
............................................................
& : member of the same patent family, corresponding
    document

EPO FORM 1503 03.82 (P04C01)